October 29, 2024

Hon. Ona T. Wang
United States Magistrate Judge
Southern District of New York
New York, New York 10007

      Re: *The New York Times Company v. Microsoft Corporation, et al.,* Case No.: 23-cv-11195-SHS-OTW *and Daily News LP, et al. v. Microsoft Corporation, et al.,* Case No.: 24-cv-03285-SHS-OTW*:* Motion to Compel Production of Documents Concerning Recently Announced and Under- Development Products

Dear Magistrate Judge Wang:

The News Plaintiffs' motion to compel about the scope of discovery into OpenAI products does not warrant Court intervention. Plaintiffs move to compel a change in product scope to broadly encompass "any current or anticipated GPT-based product (including CustomGPTs)." OpenAI has already agreed to a reasonable scope of search and production about its products, and remains willing to confer about the remainder.

To date, OpenAI has agreed to produce documents about every specific product Plaintiffs have identified: ChatGPT (including Browse (f/k/a Browse with Bing)), the OpenAI API, SearchGPT, and Microsoft's Copilot. OpenAI is also willing to produce its own core documentation about how "Custom GPTs" (which OpenAI refers to just as "GPTs") work. Furthermore, OpenAI is also making available for inspection OpenAI's models that underlie its released products. *See* Proposed Model Inspection Protocol, Dkt. No. 269. As for Plaintiffs' requests into broader, undefined product categories, the parties continue to confer about an appropriately tailored scope. For example, OpenAI agrees to search for the names of the specific third-party Custom GPTs referenced in Plaintiffs' Complaints and in their Motion, i.e. "Remove Paywall" and "News Summarizer Ace." *See* Plaintiff's Ltr. Mot., Dkt. No. 288 at 2. **Thus, rather than weigh in on this dispute now, the Court should simply require these productive conferrals to continue.**

During prior conferrals, OpenAI has asked Plaintiffs to specify what they are looking for, so that OpenAI can craft a reasonable and targeted search for the requested material. Instead of providing such specificity, the News Plaintiffs insist "the scope of this case is not limited to a narrow list of enumerated products." *Id.* Now Plaintiffs broadly seek to compel documents about "any current or anticipated GPT-based product (including CustomGPTs)." *Id.* Again, OpenAI is not saying "no." OpenAI simply does not know what Plaintiffs mean by "any current or anticipated GPT-based product (including CustomGPTs)." This is why OpenAI requests the parties continue to confer.

Custom GPTs. With respect to "current … GPT products," OpenAI understands the only remaining dispute concerns the scope of its search related to "Custom GPTs." Custom GPTs are a way for users to create "custom versions of ChatGPT that you can create for a specific purpose—called GPTs." *See generally Introducing GPTs,* OpenAI (Nov. 6, 2023),

https://openai.com/index/introducing-gpts/.   Because the platform for creating Custom GPTs on top of ChatGPT is a product itself offered by OpenAI, OpenAI is willing to produce documentation about its own platform and service. This would include, for example, OpenAI's technical and user documentation about how to create custom GPTs or search for relevant Custom GPTs in the GPT Store. As noted above, OpenAI is also willing to discuss the appropriate scope of searches to identify and produce information in its possession regarding the specific, third-party user-created GPTs referenced in the Complaint.

With respect to third-party Custom GPTs created using that product offering, OpenAI is willing to discuss with Plaintiffs a reasonable search methodology. The issue, though, is that Plaintiffs' current framing of "any current … GPT-based product," could sweep in more than three million third-party, user-created Custom GPTs. *See* Oct. 16 Agenda, Dkt. No. 262 at 27. Although Plaintiffs have recently narrowed their Custom GPT request to Custom GPTs "related to news content (including products using retrieval augmented generation to retrieve news content), product recommendations, and recipe recommendations" (Dkt. No. at 2), those broad categories of unidentified third-party customizations are still unworkable as a practical matter. Plaintiffs' proposed categories would require OpenAI to decide which third-party products Plaintiff "has in mind" as potentially relevant and make OpenAI "guess at which other [products] are encompassed by the [category]." *See Icon-IP Pty Ltd. v. Specialized Bicycle Components, Inc.*, No. 12-CV-03844-JST(MEJ), 2014 WL 4593338, at *2 (N.D. Cal. Sept. 15, 2014) (denying categorical discovery). OpenAI should not need to resort to such guessing when Plaintiffs already have access to the public list of Custom GPTs. *See* OpenAI, *Explore GPTs*, https://chatgpt.com/gpts (last visited Oct. 28, 2024); *accord Invensas Corp. v. Renesas Elecs. Corp.*, 287 F.R.D. 273, 280 (D. Del. 2012) (explaining that in ordering discovery "courts often look to . . . whether the unaccused products-at-issue are easily obtained by the plaintiff in the open market"). Plaintiffs' recent narrowing is a starting point from which the parties may be able to discuss a targeted search methodology. Again, the parties should continue to confer.

Anticipated GPT-based products. Plaintiffs have not told OpenAI what "anticipated" products they have in mind. Their motion appears to focus on the specific product SearchGPT, which is currently a limited prototype. *See generally* Dkt. No. 288 (referencing SearchGPT nearly a dozen times); *see SearchGPT Prototype*, OpenAI (July 25, 2024) https://openai.com/index/searchgpt-prototype/. Because OpenAI has now agreed to include SearchGPT in the scope of discovery, this issue is moot.

Plaintiffs' only claimed basis for discovery into *yet further* products is that "OpenAI continues to develop new products … that rely on the same basic technology at issue in the case." Dkt. No. 288 at 3. But Plaintiffs have not sought to amend or supplement their Complaints to allege that OpenAI's future products use or rely on Plaintiffs' copyrighted content, whether or not they "rely on the same basic technology."

Practically speaking, Plaintiffs' requested product definition would require OpenAI—a dynamic and rapidly-growing start-up—to supplement discovery for every contemplated or released product, regardless of whether Plaintiffs had any claim related to that product that would survive a motion to dismiss. The scope of this case has to end somewhere, and it should end at the

bounds of the claims the Plaintiffs have actually made. *See LifeNet Health v. LifeCell Corp.*, No. 2:13cv486, 2014 WL 4162113, at *6 (E.D. Va. Aug. 19, 2014) (noting that while "discovery can be used 'to assist a party to prove a claim it reasonably believes to be viable without discovery,'" it is not a method "to find out if [a party] has any basis for a claim" (citing *Micro Motion, Inc. v. Kane Steel Co.*, 894 F.2d 1318, 1327 (Fed. Cir. 1990))).

*   *   *

As shown by this opposition, OpenAI is willing to engage in good faith on a reasonable and specific discovery scope. OpenAI has already made its training data for released models available for inspection, and is ready to make the models themselves available for inspection. *See* RFP 1. OpenAI has likewise agreed to produce documents covering every specific product raised by Plaintiffs. Moreover, OpenAI has agreed to produce documents aimed at OpenAI's possession, training, referencing, or other use of the New York Times and other News Plaintiffs' content, without product limitations. *See* RFPs 74-81.[1] After the News Plaintiffs inspect those core datasets and models, if they provide a basis to revisit the topic of discovery into future products, OpenAI remains willing to confer.

Sincerely,

MORRISON FOERSTER	KEKER, VAN NEST & PETERS LLP	LATHAM & WATKINS LLP
/s/ Joseph C. Gratz	/s/ Christopher S. Sun	/s/ Elana Nightingale Dawson
Joseph C. Gratz	Christopher S. Sun	Elana Nightingale Dawson

---

[1] On Oct. 23, OpenAI agreed "to produce non-privileged documents in its possession, custody, and control (a) discussing how and why OpenAI curates and uses different types of data to train OpenAI's relevant models, subject to a reasonable search and (b) describing whether and how the relevant models were trained using articles published by The Times (if at all)." Dkt. No. 283-7 at 7 (Oct. 23 Levy Email to Frawley).